

An extremal problem with applications to testing multivariate independence

Alexander Nazarov* and Natalia Stepanova†

Abstract

Some problems of statistics can be reduced to extremal problems of minimizing functionals of smooth functions defined on the cube $[0, 1]^m$, $m \geq 2$. In this paper, we study a class of extremal problems that is closely connected to the problem of testing multivariate independence. By solving the extremal problem, we provide a unified approach to establishing weak convergence for a wide class of empirical processes which emerge in connection with testing independence. The use of our result is also illustrated by describing the domain of local asymptotic optimality of some nonparametric tests of independence.

Keywords: boundary-value problem; Green function; multivariate independence; asymptotic efficiency of test statistics; local asymptotic optimality

AMS Subject Classification: 62G10; 62G20; 35C05

1 Introduction

Let $\mathbf{X}_i = (X_{i1}, \dots, X_{im})$, $m \geq 2$, $i = 1, \dots, n$, be independent random vectors with absolutely continuous distribution function (df) F and marginal df's F_1, \dots, F_m . An important problem of statistics is to test the hypothesis of (multivariate) independence

$$H_0 : F \equiv F_1 F_2 \dots F_m. \quad (1)$$

Many test statistics for testing multivariate independence converge weakly toward functionals of Gaussian random processes under the hypothesis of independence (see, e.g., Deheuvels (1981); Nikitin (1995); Nazarov and Nikitin (2000); Schmid and Schmidt (2007)). If such functionals are regular then the limiting distributions of test statistics are easily derived. The knowledge of the limiting distribution is important for calculating (approximate) critical values and finding

**Department of Mathematics and Mechanics, St. Petersburg State University, St. Petersburg, 198504, Russia*

†*School of Mathematics and Statistics, Carleton University, Ottawa, ON, K1S 5B6 Canada*

asymptotic efficiency of the sequence of test statistics. Unfortunately, many independence tests have a complex structure, and the corresponding asymptotic theory is not easy to develop. This prevents the use of such tests. In this paper, by solving a multidimensional extremal problem (Theorem 1), we prove a result on the weak convergence of empirical processes (Theorem 2), which yields the limit distribution for a wide class of test statistics for testing independence. Such a class includes, as extreme cases, the multivariate versions of Cramér–von Mises and Blum–Kiefer–Rosenblatt test statistics (see Section 4.1). In addition, we illustrate the use of Theorem 1 for describing the domain of local asymptotic optimality of some nonparametric tests of independence (see Section 4.2).

Let $M = \{1, 2, \dots, m\}$ and let 2^M be the set of all subsets of M . In this paper, we study a class of extremal problems indexed by a subset \mathcal{M} of 2^M that obeys certain restrictions. For $m = 2$ there exist three essentially different choices for \mathcal{M} , and the class of extremal problems is limited. Such problems were studied in (Nikitin, 1995, Ch. 5) and Nazarov and Nikitin (2000) in connection with calculating asymptotic efficiency of nonparametric tests of independence in a bivariate setup. In a general case of $m \geq 2$ the situation is more complicated, and the number of possible extremal problems for an arbitrary m is not explicitly calculable (see Remark 1).

The paper is organized as follows. In Sections 2 and 3 we formulate and solve the extremal problem of interest. The solution is obtained by reducing the extremal problem to a non-standard boundary-value problem and constructing a Green function for the latter. This part of the paper, with the main result given by Theorem 1, is purely analytical and might be of independent interest for experts in PDE's. The rest of the paper is largely statistical, with a focus on various applications of Theorem 1 to the problem of testing independence. In particular, Theorem 1 provides a unified approach to establishing weak convergence for a wide class of empirical processes with a multidimensional time parameter which emerge in nonparametric statistics (Theorem 2).

2 Extremal problem

Let $I^m = [0, 1]^m$ and let \mathbf{C}_0^m be the space of (real-valued) functions that are m -times continuously differentiable with respect to each variable and obey the following boundary conditions:

$$\mathbf{C}_0^m = \{\Omega \in \mathbf{C}^m(I^m) : \Omega(\mathbf{x})|_{x_j=0} = 0, \quad j = 1, \dots, m\},$$

where $\mathbf{x} = (x_1, \dots, x_m) \in I^m$. Define a scalar product on \mathbf{C}_0^m as follows:

$$(\Omega_1, \Omega_2) = \int_{I^m} \omega_1(\mathbf{x}) \omega_2(\mathbf{x}) d\mathbf{x}, \quad \Omega_1, \Omega_2 \in \mathbf{C}_0^m, \quad (2)$$

where $\omega_i(\mathbf{x}) = \frac{\partial^m \Omega_i(\mathbf{x})}{\partial x_1 \dots \partial x_m}$, $i = 1, 2$. Denote by \mathbf{H}^m the closure of the space \mathbf{C}_0^m under the norm $\|\cdot\|$ induced by the scalar product (2). For any $m \geq 2$, \mathbf{H}^m is a

Hilbert space whose properties are derived similarly to the case $m = 2$ studied in Nazarov and Nikitin (2000). More precisely, the following result holds true.

Proposition 1 (a) *The embedding of the space \mathbf{H}^m into the space $\mathbf{C}(I^m)$ is compact.* (b) *The embedding of the space \mathbf{H}^m into the Sobolev space $\mathbf{W}_2^1(I^m)$ is compact.*

Part (a) of Proposition 1 implies that any function from \mathbf{H}^m equals zero on any “left” side of the cube I^m adjacent to the origin.

Consider the extremal problem

$$\|\Omega\|_{\mathbf{H}^m} \rightarrow \min, \quad (3)$$

$$\Omega \in \mathbf{H}_{\mathcal{M}}^m, \quad \int_{I^m} \Omega(\mathbf{x}) d\mu(\mathbf{x}) = 1, \quad (4)$$

where $\mathbf{H}_{\mathcal{M}}^m$ is a subset of \mathbf{H}^m specified by certain boundary conditions on the “right” sides of the cube I^m adjacent to the point $\mathbf{1} = (1, 1, \dots, 1)$, and μ is a finite measure on I^m . In order to describe possible boundary conditions of this extremal problem we need some notation.

Let $M = \{1, 2, \dots, m\}$ and let 2^M be the set of all subsets of M . For any $U \subset M$, denote \mathbf{x}_U the $|U|$ -dimensional vector $\mathbf{x}_U = (x_i : i \in U)$, where $|U|$ is the cardinality of U . Consider the set $\mathcal{M} \subset 2^M$ such that

$$\forall U \subset V \subset 2^M, \quad U \in \mathcal{M} \text{ implies } V \in \mathcal{M}. \quad (5)$$

That is, if set U belongs to \mathcal{M} , then all its “oversets” also belong to \mathcal{M} . Define the subset $\mathbf{H}_{\mathcal{M}}^m$ of \mathbf{H}^m as follows:

$$\mathbf{H}_{\mathcal{M}}^m = \{\Omega \in \mathbf{H}^m : \Omega(\mathbf{x})|_{\mathbf{x}_U=\mathbf{1}} = 0, \quad U \in \mathcal{M}\}.$$

The reason for the requirement (5) is simple: if $\Omega \in \mathbf{H}^m$ takes a zero value on the side $S_U = \{\mathbf{x}_U = \mathbf{1}\}$, it also takes a zero value on all the subedges of S_U of less dimension.

For a set $V = (i_1, \dots, i_l)$ and its complement (in M) $V^c = (j_1, \dots, j_k)$, $l+k = m$, put

$$\partial_{\mathbf{x}_V} \partial_{\mathbf{x}_{V^c}}^2 = \partial_{x_{i_1}} \dots \partial_{x_{i_l}} \partial_{x_{j_1}}^2 \dots \partial_{x_{j_k}}^2.$$

By the Lagrange principle rule (see, e.g., Alexeev et al. (1979)), the necessary condition of a minimum in (3)–(4) is reduced to the Euler equation (in the sense of distributions)

$$(-1)^m \lambda \frac{\partial^{2m} \Omega(\mathbf{x})}{\partial x_1^2 \dots \partial x_m^2} = \mu(\mathbf{x}), \quad (6)$$

where λ is the Lagrange multiplier, and the natural boundary conditions

$$\left. \frac{\partial^{l+2k} \Omega(\mathbf{x})}{\partial_{\mathbf{x}_V} \partial_{\mathbf{x}_{V^c}}^2} \right|_{\mathbf{x}_V=\mathbf{1}} = 0, \quad \text{for any } V \notin \mathcal{M}, \quad V \neq \emptyset. \quad (7)$$

Remark 1 For any $U \in 2^M$ define an m -dimensional vector of Boolean variables ($y_j = \mathbb{I}(j \in U), j = 1, \dots, m$). Then $\mathbb{I}(U \in \mathcal{M})$ is a monotone Boolean function (see Kleitman (1969)). Denote by $N(m)$ the total number of such functions. Obviously, the number of the above considered extremal problems is also equal to $N(m)$. So far, no explicit formula for $N(m)$ as a function of m has been found. For the asymptotic behavior of $N(m)$, as $m \rightarrow \infty$, see Korshunov (1981).

3 The Green function of the problem

Recall that Green function of the boundary-value problem (4), (6), (7) is the function $\mathcal{G}_{\mathcal{M}}(\mathbf{x}, \boldsymbol{\xi})$ that satisfies, along with the boundary conditions, the equation (in the sense of distributions)

$$(-1)^m \frac{\partial^{2m} \mathcal{G}_{\mathcal{M}}(\mathbf{x}, \boldsymbol{\xi})}{\partial x_1^2 \dots \partial x_m^2} = \delta(\mathbf{x} - \boldsymbol{\xi}), \quad (8)$$

where $\delta(\mathbf{x})$ is the Dirac function. It is well known that solution to the problem (4), (6), (7), and hence solution to the extremal problem (3)–(4), can be expressed with the aid of Green function by the formula

$$\Omega(\mathbf{x}) = \lambda^{-1} \int_{I^m} \mathcal{G}_{\mathcal{M}}(\mathbf{x}, \boldsymbol{\xi}) d\mu(\boldsymbol{\xi}), \quad \mathbf{x} \in I^m, \quad (9)$$

where the Lagrange multiplier λ is found from the integral restriction in (4) and has the form

$$\lambda = \iint_{I^m \times I^m} \mathcal{G}_{\mathcal{M}}(\mathbf{x}, \boldsymbol{\xi}) d\mu(\mathbf{x}) d\mu(\boldsymbol{\xi}). \quad (10)$$

For $\mathbf{x}, \boldsymbol{\xi} \in I^m$ and for $V \subset M$ as before, define the functions

$$k_V(\mathbf{x}, \boldsymbol{\xi}) = k_{i_1}(\mathbf{x}, \boldsymbol{\xi}) \dots k_{i_l}(\mathbf{x}, \boldsymbol{\xi}), \quad K_{V^c}(\mathbf{x}, \boldsymbol{\xi}) = K_{j_1}(\mathbf{x}, \boldsymbol{\xi}) \dots K_{j_k}(\mathbf{x}, \boldsymbol{\xi}),$$

where

$$k_j(\mathbf{x}, \boldsymbol{\xi}) = x_j \xi_j, \quad K_j(\mathbf{x}, \boldsymbol{\xi}) = \min(x_j, \xi_j), \quad j = 1, \dots, m. \quad (11)$$

The first main result of the paper is stated as follows.

Theorem 1 *The Green function of the boundary-value problem (4), (6), (7) is*

$$\mathcal{G}_{\mathcal{M}}(\mathbf{x}, \boldsymbol{\xi}) = K_M(\mathbf{x}, \boldsymbol{\xi}) - \sum_{U \in \mathcal{M}} a_U K_{U^c}(\mathbf{x}, \boldsymbol{\xi}) k_U(\mathbf{x}, \boldsymbol{\xi}), \quad (12)$$

with the coefficients a_U defined recurrently by

$$\sum_{\substack{V \subset U \\ V \in \mathcal{M}}} a_V = 1, \quad \text{for all } U \in \mathcal{M}. \quad (13)$$

Proof First, note that

$$-\frac{\partial^2 K_j(\mathbf{x}, \boldsymbol{\xi})}{\partial x_j^2} = \delta(x_j - \xi_j), \quad \frac{\partial^2 k_j(\mathbf{x}, \boldsymbol{\xi})}{\partial x_j^2} = 0.$$

Therefore the function $\mathcal{G}_{\mathcal{M}}$ satisfies (8) with an arbitrary choice of the constants a_U .

Further, by (5) for any nonempty set $V = (i_1, \dots, i_l) \notin \mathcal{M}$ and for any $U \in \mathcal{M}$ there exists $j \in U^c \cap V$, and hence

$$\frac{\partial^l K_{U^c}(\mathbf{x}, \boldsymbol{\xi}) k_U(\mathbf{x}, \boldsymbol{\xi})}{\partial \mathbf{x}_V} \Big|_{\mathbf{x}_V = \mathbf{1}} = \frac{\partial K_j(\mathbf{x}, \boldsymbol{\xi})}{\partial x_j} \Big|_{x_j = 1} \times \dots = 0.$$

Thus all summands in (12) satisfy (7).

In order to prove that the function $\mathcal{G}_{\mathcal{M}}$ vanishes on $\{\mathbf{x}_U = \mathbf{1}\}$ for some $U \in \mathcal{M}$, we represent \mathcal{M} as the union $\mathcal{M} = \bigcup_{V \supset U} \mathfrak{M}_V$ of the disjoint sets

$$\mathfrak{M}_V = \{W \in \mathcal{M} : W \cup U = V\}$$

(in the definition of \mathfrak{M}_V the basic property (5) is used).

The key observation is that for any $W \in \mathfrak{M}_V$,

$$K_{W^c}(\mathbf{x}, \boldsymbol{\xi}) k_W(\mathbf{x}, \boldsymbol{\xi}) \Big|_{\mathbf{x}_U = \mathbf{1}} = K_{V^c}(\mathbf{x}, \boldsymbol{\xi}) k_V(\mathbf{x}, \boldsymbol{\xi}) \Big|_{\mathbf{x}_U = \mathbf{1}}.$$

Therefore formula (12) gives

$$\begin{aligned} \mathcal{G}_{\mathcal{M}}(\mathbf{x}, \boldsymbol{\xi}) \Big|_{\mathbf{x}_U = \mathbf{1}} &= K_{U^c}(\mathbf{x}, \boldsymbol{\xi}) k_U(\mathbf{x}, \boldsymbol{\xi}) \Big|_{\mathbf{x}_U = \mathbf{1}} \cdot \left(1 - \sum_{W \in \mathfrak{M}_U} a_W\right) - \\ &\quad - \sum_{\substack{V \supset U \\ V \neq U}} K_{V^c}(\mathbf{x}, \boldsymbol{\xi}) k_V(\mathbf{x}, \boldsymbol{\xi}) \Big|_{\mathbf{x}_U = \mathbf{1}} \cdot \sum_{W \in \mathfrak{M}_V} a_W. \end{aligned}$$

Due to (13) the coefficients on the right-hand side vanish, and $\mathcal{G}_{\mathcal{M}}(\mathbf{x}, \boldsymbol{\xi}) \Big|_{\mathbf{x}_U = \mathbf{1}} = 0$. This completes the proof. \square

Remark 2 Return to the extremal problem (3)–(4) and consider the following three sets of boundary conditions: (a) there are no restrictions on $\Omega \in \mathbf{H}^m$ except for those that specify the space \mathbf{H}^m , (b) $\Omega \in \mathbf{H}^m$ equals zero on any $(m-1)$ -dimensional side of I^m , and (c) $\Omega \in \mathbf{H}^m$ equals zero at the point $\mathbf{1} = (1, \dots, 1)$. Then $\mathcal{M} = \emptyset$, $\mathcal{M} = 2^M$, $\mathcal{M} = \{M\}$, respectively, and by Theorem 1 the corresponding Green functions are $\prod_{j=1}^m K_j(\mathbf{x}, \boldsymbol{\xi})$, $\prod_{j=1}^m (K_j(\mathbf{x}, \boldsymbol{\xi}) - k_j(\mathbf{x}, \boldsymbol{\xi}))$, $\prod_{j=1}^m K_j(\mathbf{x}, \boldsymbol{\xi}) - \prod_{j=1}^m k_j(\mathbf{x}, \boldsymbol{\xi})$. These are covariance functions of the classical Gaussian random processes. They correspond to an m -dimensional Brownian sheet, an m -dimensional Brownian pillow, and an m -dimensional tucked Brownian sheet, respectively. The latter two arise as limiting processes in nonparametric testing of independence (see Section 4.1 for details).

4 Connection to testing independence

This section illustrates the use of Theorem 1 for some efficiency issues that emerge in the problem of testing multivariate independence. We use a general dependence model which is rather popular in the present context.

Let $\mathbf{X}_i = (X_{i1}, \dots, X_{im})$, $m \geq 2$, $i = 1, \dots, n$, be independent random vectors with absolutely continuous df F and marginal df's F_1, \dots, F_m . Consider the problem of testing the hypothesis of independence (1).

For the absolutely continuous df F , a *copula* ϕ of F is a (unique) df with uniform univariate margins, such that $\phi(F_1(x_1), \dots, F_m(x_m)) = F(x_1, \dots, x_m)$. When testing for independence using distribution-free (independent of F_1, \dots, F_m) test statistics, one can assume uniform margins and define the model in terms of copulas.

Let $\{F_\theta : \theta \geq 0\}$ be the family of absolutely continuous copulas F_θ that are monotone in θ (the case $\theta = 0$ corresponds to independence) and that satisfy the following common regularity conditions, cf. Genest et al. (2007); Quesy (2009):

- (A1) the density $f_\theta(\mathbf{x}) = \partial^m F_\theta(\mathbf{x}) / \partial x_1 \dots \partial x_m$ admits a square-integrable derivative $\dot{f}_0(\mathbf{x})$ of $f_\theta(\mathbf{x})$ with respect to θ at $\theta = 0$ for every $\mathbf{x} = (x_1, \dots, x_m) \in I^m$, and the function $\sqrt{f_\theta(\mathbf{x})}$ is differentiable in quadratic mean at $\theta = 0$, i.e.,

$$\int_{I^m} \left(\sqrt{f_\theta(\mathbf{x})} - 1 - \frac{1}{2} \theta \dot{f}_0(\mathbf{x}) \right)^2 d\mathbf{x} = o(\theta^2), \quad \theta \rightarrow 0,$$

- (A2) $F_\theta(\mathbf{x})$ is differentiable with respect to θ in a small neighborhood of $\theta = 0$ for every $\mathbf{x} \in I^m$, and the following identity holds for every $\mathbf{x} \in I^m$:

$$\dot{F}_0(\mathbf{x}) = \lim_{\theta \rightarrow 0} \frac{\partial}{\partial \theta} F_\theta(\mathbf{x}) = \int_0^{x_1} \dots \int_0^{x_m} \dot{f}_0(y_1, \dots, y_m) dy_1 \dots dy_m;$$

and, in addition,

$$\int_{I^m} \dot{F}_0(\mathbf{x}) d\mathbf{x} = \lim_{\theta \rightarrow 0} \frac{\partial}{\partial \theta} \int_{I^m} F_\theta(\mathbf{x}) d\mathbf{x}.$$

The function $\dot{F}_0(\mathbf{x})$ is sometimes called the *dependence function*. Assumption (A2) implies that $F_\theta(\mathbf{x})$ can be written in terms of \dot{F}_0 as follows:

$$F_\theta(\mathbf{x}) = \prod_{j=1}^m x_j + \theta \dot{F}_0(\mathbf{x}) + o(\theta), \quad \mathbf{x} \in I^m, \quad \theta \rightarrow 0,$$

where, by the property of a multivariate copula, the boundary conditions

$$\dot{F}_0(\mathbf{x})|_{x_k=0} = 0, \quad \dot{F}_0(\mathbf{x})|_{x_U=1} = 0, \quad \text{for any } U \in 2^M, \quad |U| = m - 1 \quad (14)$$

are satisfied. As an alternative to $H_0 : \theta = 0$ we consider the hypothesis $H_1 : \theta > 0$. In what follows, the underlying df is assumed to belong to the family $\{F_\theta : \theta \geq 0\}$.

The family $\{F_\theta : \theta \geq 0\}$ produces a sequence of locally asymptotically normal experiments. Indeed, let \mathbf{P}_θ be the probability distribution used to calculate the df F_θ . Then the full observation is a single observation from the product \mathbf{P}_θ^n of n copies of \mathbf{P}_θ . In view of condition **(A1)**, the sequence of statistical experiments $\{\mathbf{P}_{h/\sqrt{n}}^n : h \geq 0\}$, indexed by a local parameter $h = \sqrt{n}\theta$, is locally asymptotically normal (LAN) at the point $h = 0$, that is,

$$\log \frac{d\mathbf{P}_{h/\sqrt{n}}^n}{d\mathbf{P}_0^n}(\mathbf{X}_1, \dots, \mathbf{X}_n) = h\Delta_{n,0} - \frac{1}{2}h^2 I_0 + o_{\mathbf{P}_0^n}(1), \quad n \rightarrow \infty,$$

where $I_\theta = \int_{I^m} \dot{f}_\theta^2(\mathbf{x})/f_\theta(\mathbf{x}) d\mathbf{x}$ is the Fisher information in the family $\{F_\theta, \theta \geq 0\}$ and $\Delta_{n,0} = n^{-1/2} \sum_{i=1}^n \dot{f}_0(\mathbf{X}_i) \overset{\mathbf{P}_0^n}{\rightsquigarrow} N(0, I_0)$. The symbol $\overset{\mathbf{P}_0^n}{\rightsquigarrow}$ denotes convergence in \mathbf{P}_0^n -distribution. The LAN property ensures the mutual contiguity of the sequences of distributions $\{\mathbf{P}_{h/\sqrt{n}}^n\}$ and $\{\mathbf{P}_0^n\}$, and facilitates the calculation of Pitman efficiency of asymptotically normal test statistics (see (van der Vaart, 1998, Ths. 14.7, 15.4)).

4.1 Asymptotic efficiency of independence tests

Consider testing the hypothesis of independence using distribution-free statistics of $\mathbf{X}_1, \dots, \mathbf{X}_n$ in the case when some of the F_j 's, $j = 1, \dots, m$, are *known*, while the others are *unknown*. Denote by \mathbb{F}_n the multivariate empirical df that corresponds to F , and denote by $\mathbb{F}_{j,n}$ the marginal empirical df based on X_{1j}, \dots, X_{nj} , $j = 1, \dots, m$. For a set $V = (i_1, \dots, i_l)$ and its complement (in M) $V^c = (j_1, \dots, j_k)$, $l + k = m$, put

$$\begin{aligned} F_V(\mathbf{x}) &= F_{i_1}(x_{i_1}) \dots F_{i_l}(x_{i_l}), & dF_V(\mathbf{x}) &= dF_{i_1}(x_{i_1}) \dots dF_{i_l}(x_{i_l}), \\ \mathbb{F}_{V^c,n}(\mathbf{x}) &= \mathbb{F}_{j_1,n}(x_{j_1}) \dots \mathbb{F}_{j_k,n}(x_{j_k}), & d\mathbb{F}_{V^c,n}(\mathbf{x}) &= d\mathbb{F}_{j_1,n}(x_{j_1}) \dots d\mathbb{F}_{j_k,n}(x_{j_k}). \end{aligned}$$

For $p = 1, 2, \dots$, consider the class $\{B_{V,n}^p : V \subset M\}$ of test statistics, cf. Dugué (1975),

$$B_{V,n}^p = \int_{\mathbb{R}^m} (\mathbb{F}_n(\mathbf{x}) - F_V(\mathbf{x})\mathbb{F}_{V^c,n}(\mathbf{x}))^p dF_V(\mathbf{x}) d\mathbb{F}_{V^c,n}(\mathbf{x}).$$

Suppose that the margins F_j , $j \in V$, are known. Then the test statistics $B_{V,n}^p$, $V \subset M$, $p = 1, 2, \dots$, are distribution-free under the null hypothesis, and the study of their behaviour under H_0 can be done when $F(\mathbf{x})$ is a uniform distribution on I^m , which will be assumed from now on. Choosing $V = M$ yields the Cramér–von Mises-type statistics

$$B_{M,n}^p = \int_{I^m} (\mathbb{F}_n(\mathbf{x}) - \mathbf{x}_M)^p d\mathbf{x}_M,$$

where $\mathbf{x}_M = x_1 \dots x_m$ and $d\mathbf{x}_M = d\mathbf{x} = dx_1 \dots dx_m$. On the other hand, setting $V = \emptyset$ leads to the Blum–Kiefer–Rosenblatt-type statistics

$$B_{\emptyset,n}^p = \int_{I^m} (\mathbb{F}_n(\mathbf{x}) - \mathbb{F}_{M,n}(\mathbf{x}))^p d\mathbb{F}_{M,n}(\mathbf{x}).$$

An important step in calculating asymptotic efficiency of the sequence of test statistics $\{B_{V,n}^p\}_{n \geq 1}$ lies in showing the weak convergence of

$$W_{V,n}(\mathbf{x}) = \sqrt{n}(\mathbb{F}_n(\mathbf{x}) - \mathbf{x}_V \mathbb{F}_{V^c,n}(\mathbf{x})), \quad \mathbf{x} \in I^m,$$

to a limiting Gaussian process under the null hypothesis. This is achieved with the aid of Theorem 1, and is stated below as Theorem 2. The Skorohod space $D(I^m)$ that appears in the statement of Theorem 2 and generalizes the well-known space $D[0, 1]$, is described in Neuhaus (1971), where some of its properties are also derived.

Theorem 2 *Assume that $F(\mathbf{x})$ is a uniform distribution on I^m . Then for any $V \subset M$ the empirical process $W_{V,n}(\mathbf{x})$ converges weakly in the Skorohod space $D(I^m)$ to a centered Gaussian process $W_V(\mathbf{x})$ with covariance function $\mathcal{G}_{\mathcal{M}_V}(\mathbf{x}, \boldsymbol{\xi})$ given by (12), where the set $\mathcal{M}_V \subset 2^M$ is defined as follows:*

$$\mathcal{M}_V = \{M, M \setminus U : U \subset V^c, |U| = 1\}. \quad (15)$$

Proof The proof follows the pattern of (Neuhaus, 1971, Sec. 4, 5), where, among others, weak convergence of the process $W_{M,n}(\mathbf{x}) = \sqrt{n}(\mathbb{F}_n(\mathbf{x}) - \mathbf{x}_M)$ to a tucked Brownian sheet is established. Therefore, most details on the convergence of finite-dimensional distributions of $W_{V,n}(\mathbf{x})$ to a multivariate normal distribution, and the proof of tightness of the family of associated probability measures are omitted. A pivotal point of our proof is obtaining an expression for the covariance of the limiting process. This is accomplished by appealing to Theorem 1.

For some $j \notin V$ let $U = M \setminus \{j\}$. Then $\mathbb{F}_n(\mathbf{x})|_{x_U=\mathbf{1}} = \mathbb{F}_{j,n}(x_j)$, so that the process $W_{V,n}(\mathbf{x})$ is pinned down to zero on $\{\mathbf{x}_U = \mathbf{1}\}$ for any $U \in \mathcal{M}$ such that $V \subset U$ and $|U| = m - 1$.

Now consider the boundary-value problem (4), (6), (7) with the set $\mathcal{M} = \mathcal{M}_V$ given by (15). The respective boundary condition takes the form

$$\Omega \in \mathbf{H}^m, \quad \Omega(\mathbf{x})|_{\mathbf{x}_U=\mathbf{1}} = 0 \quad \text{for any } U \in \mathcal{M} \text{ such that } V \subset U, |U| = m - 1.$$

The function $\Omega(\mathbf{x})$ equals zero exactly on those sides $\{\mathbf{x}_U = \mathbf{1}\}$ of the cube I^m where the empirical process $W_{V,n}(\mathbf{x})$ vanishes. This observation together with Theorem 1 gives the required expression for the covariance function. The proof is completed. \square

In the two extreme cases, when (1) $V = M$ and (2) $V = \emptyset$, Theorem 2 yields a well-known result (see, e.g., Blum et al. (1961); Neuhaus (1971); Deheuvels (1981)). Indeed, in the first case the set in (15) reduces to $\mathcal{M}_M = \{M\}$ and the covariance function of $W_M(\mathbf{x})$ is

$$\mathcal{G}_{\mathcal{M}_M}(\mathbf{x}, \boldsymbol{\xi}) = \prod_{j=1}^m K_j(\mathbf{x}, \boldsymbol{\xi}) - \prod_{j=1}^m k_j(\mathbf{x}, \boldsymbol{\xi}), \quad (16)$$

The respective boundary condition has the form $\Omega \in \mathbf{H}^m$, $\Omega(\mathbf{x})|_{\mathbf{x}_M=1} = 0$. In the second case (15) becomes

$$\mathcal{M}_\emptyset = \{M, M \setminus \{1\}, M \setminus \{2\}, \dots, M \setminus \{m\}\},$$

and the covariance function of $W_\emptyset(\mathbf{x})$ is

$$\mathcal{G}_{\mathcal{M}_\emptyset}(\mathbf{x}, \boldsymbol{\xi}) = \prod_{j=1}^m K_j(\mathbf{x}, \boldsymbol{\xi}) - \sum_{j=1}^m K_j(\mathbf{x}, \boldsymbol{\xi}) \prod_{i \neq j} k_i(\mathbf{x}, \boldsymbol{\xi}) + (m-1) \prod_{j=1}^m k_j(\mathbf{x}, \boldsymbol{\xi}). \quad (17)$$

The respective boundary condition takes the form $\Omega \in \mathbf{H}^m$, $\Omega(\mathbf{x})|_{\mathbf{x}_U=1} = 0$ for any $U \in 2^M$ such that $|U| = m-1$.

For large sample sizes, the quality of test statistics $B_{V,n}^p$, $V \subset M$, $p \geq 1$, can be judged by looking at their Bahadur efficiency. This kind of asymptotic efficiency is quantified by the *Bahadur exact slope*. Finding the Bahadur exact slope of a sequence of test statistics requires the law of large numbers under the alternative, and the rough large deviation asymptotics under the null hypothesis. The problem of large deviation asymptotics consists in minimizing the Kullback–Leibler information over a subset of distribution functions that depends on the structure of the test statistic. When $p = 1$ the above minimization problem is reduced, by using variational methods, to the boundary-value problem (4), (6), (7) with μ being the Lebesgue measure on I^m and $\mathcal{M} = \mathcal{M}_V$ given by (15), whose solution provides the main contribution to the initial problem (see (Nikitin, 1995, Ch. 5) for details).

For example, finding the rough large deviation asymptotics of the statistic

$$B_{\emptyset,n}^1 = \int_{I^m} (\mathbb{F}_n(\mathbf{x}) - \mathbb{F}_{M,n}(\mathbf{x})) d\mathbb{F}_{M,n}(\mathbf{x})$$

is largely reduced to the boundary-value problem (4), (6), (7) with $\mathcal{M} = \mathcal{M}_\emptyset$ and for sufficiently small $t > 0$, cf. formula (5.3.29) of Nikitin (1995),

$$\lim_{n \rightarrow \infty} n^{-1} \log \mathbf{P}_{H_0}(B_{\emptyset,n}^1 \geq t) = -\frac{1}{2} \lambda_0 t^2 + \sum_{j \geq 3} c_j t^j, \quad (18)$$

where the series on the right-hand side is convergent and, cf. (9) and (10),

$$\lambda_0 = \left(\iint_{I^m \times I^m} \mathcal{G}_{\mathcal{M}_\emptyset}(\mathbf{x}, \boldsymbol{\xi}) d\mathbf{x} d\boldsymbol{\xi} \right)^{-1} = \frac{4^m}{(4/3)^m - m/3 - 1}$$

(see (Nikitin, 1995, Sec. 5.3)). Since, under the alternative,

$$B_{\emptyset,n}^1 \xrightarrow{\mathbf{P}_\theta} \theta \int_{I^m} \dot{F}_0(\mathbf{x}) d\mathbf{x}, \quad n \rightarrow \infty,$$

where the symbol $\xrightarrow{\mathbf{P}_0^n}$ denotes convergence in \mathbf{P}_0^n -probability, it follows from (18) and Theorem 1.2.2 of Nikitin (1995) that the Bahadur exact slope $c_{B_\emptyset^1}(\theta)$ of the sequence $\{B_{\emptyset,n}^1\}_{n \geq 1}$ satisfies as $\theta \rightarrow 0$

$$c_{B_\emptyset^1}(\theta) \sim \theta^2 \frac{4^m}{(4/3)^m - m/3 - 1} \left(\int_{I^m} \dot{F}_0(\mathbf{x}) d\mathbf{x} \right)^2. \quad (19)$$

Similarly, in a general case, for the model under consideration a routine computation leads to the following result.

Proposition 2 *For an arbitrary $V \subset M$ the Bahadur exact slope of the sequence $\{B_{V,n}^1\}_{n \geq 1}$ satisfies*

$$c_{B_V^1}(\theta) \sim \theta^2 \left(\iint_{I^m \times I^m} \mathcal{G}_{\mathcal{M}_V}(\mathbf{x}, \boldsymbol{\xi}) d\mathbf{x} d\boldsymbol{\xi} \right)^{-1} \left(\int_{I^m} \dot{F}_0(\mathbf{x}) d\mathbf{x} \right)^2, \quad \theta \rightarrow 0,$$

where the function $\mathcal{G}_{\mathcal{M}_V}(\mathbf{x}, \boldsymbol{\xi})$ is the same as in Theorem 2.

Compared to $B_{V,n}^1$, the evaluation of the local Bahadur efficiency of the tests based on $B_{V,n}^p$, $p \geq 2$, is much more complicated. For example, when $p = 2$, the efficiency problem is reduced to calculating the principal eigenvalue of the integral operator with kernel $\mathcal{G}_{\mathcal{M}_V}$ (see (Nikitin, 1995, Ch. 5) for details).

A more complex empirical process that vanishes completely at the boundary of I^m was studied, for example, in Neuhaus (1971) and Deheuvels (2005). Such a process converges weakly to the m -dimensional Brownian pillow with covariance function

$$\mathcal{G}_{2^M}(\mathbf{x}, \boldsymbol{\xi}) = \prod_{j=1}^m (K_j(\mathbf{x}, \boldsymbol{\xi}) - k_j(\mathbf{x}, \boldsymbol{\xi})), \quad \mathbf{x}, \boldsymbol{\xi} \in I^m, \quad (20)$$

which corresponds to the Green function (12) with $\mathcal{M} = 2^M$ (see Remark 2), and appears in connection with testing multivariate independence in the following context.

Consider $W_{\emptyset,n}(\mathbf{x}) = \sqrt{n}(\mathbb{F}_n(\mathbf{x}) - \prod_{j=1}^m \mathbb{F}_{j,n}(x_j))$. The corresponding limiting process $W_\emptyset(\mathbf{x})$ has the covariance function $\mathcal{G}_{\mathcal{M}_\emptyset}(\mathbf{x}, \boldsymbol{\xi})$ which coincides with $\mathcal{G}_{2^M}(\mathbf{x}, \boldsymbol{\xi})$ when $m = 2$. However, for $m \geq 3$ the situation changes. The process $W_\emptyset(\mathbf{x})$ does not vanish completely at the facets of I^m adjacent to the point $\mathbf{x} = \mathbf{1}$, but only does so at the one-dimensional edges (see (17)). This “disappointing” property is overcome by the *tied-down* empirical process (see, e.g., Neuhaus (1971); Deheuvels (2005))

$$\hat{W}_{\emptyset,n}(\mathbf{x}) = \sqrt{n} \left(\mathbb{F}_n(\mathbf{x}) - \sum_{k=1}^m (-1)^{k-1} \sum_{U \subset M: |U|=k} \mathbf{x}_U \cdot \mathbb{F}_n(\mathbf{x})|_{\mathbf{x}_U=1} \right), \quad \mathbf{x} \in I^m,$$

which can be equivalently written in the form

$$\hat{W}_{\emptyset,n}(\mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \prod_{j=1}^m (\mathbb{I}(X_{ij} \leq x_j) - x_j), \quad \mathbf{x} \in I^m.$$

Under the null hypothesis, $\hat{W}_{\emptyset,n}(\mathbf{x})$ converges weakly in the Skorokhod space $D(I^m)$ toward an m -dimensional Brownian pillow $\hat{W}_{\emptyset}(\mathbf{x})$ with covariance function (20). This fact seems to be established for the first time in Neuhaus (1971). The corresponding test statistics (for $m \geq 3$) take the form, cf. Dugué (1975); Deheuvels (1981),

$$\hat{B}_{m,n}^p = \int_{I^m} \left(\mathbb{F}_n(\mathbf{x}) - \sum_{k=1}^m (-1)^{k-1} \sum_{U \subset M: |U|=k} \mathbf{x}_U \cdot \mathbb{F}_n(\mathbf{x})|_{\mathbf{x}_U=\mathbf{1}} \right)^p d\mathbf{x}.$$

For $p = 1$, under the hypothesis of independence, the limiting distribution of $\sqrt{n}\hat{B}_{m,n}^1$ is normal with zero mean and variance, cf. Th. 6 of Deheuvels (1981),

$$\hat{\sigma}_m^2(0) = \iint_{I^m \times I^m} \mathcal{G}_{2^M}(\mathbf{x}, \boldsymbol{\xi}) d\mathbf{x} d\boldsymbol{\xi} = 12^{-m}.$$

Therefore, by Le Cam's third lemma, for all $h \geq 0$,

$$\frac{\sqrt{n}(\hat{B}_{m,n}^1 - \hat{\mu}_m(h/\sqrt{n}))}{\hat{\sigma}_m(0)} \overset{\mathbf{P}_{h/\sqrt{n}}^n}{\rightsquigarrow} \mathcal{N}(0, 1), \quad n \rightarrow \infty,$$

where, using (14),

$$\hat{\mu}_m(\theta) = \theta \int_{I^m} \left(\dot{F}_0(\mathbf{x}) - \sum_{k=1}^{m-2} (-1)^{k-1} \sum_{U \subset M: |U|=k} \mathbf{x}_U \cdot \dot{F}_0(\mathbf{x})|_{\mathbf{x}_U=\mathbf{1}} \right) d\mathbf{x}$$

The symbol $\overset{\mathbf{P}_{\theta}^n}{\rightsquigarrow}$ denotes convergence in \mathbf{P}_{θ}^n -distribution of a random sample drawn from F_{θ} . In view of Theorem 14.7 in van der Vaart (1998), the squared Pitman slope of the sequence $\{\hat{B}_{m,n}^1\}_{n \geq 1}$ is

$$\left(\frac{\hat{\mu}'_m(0)}{\hat{\sigma}_m(0)} \right)^2 = 12^m \left(\int_{I^m} \left(\dot{F}_0(\mathbf{x}) - \sum_{k=1}^{m-2} (-1)^{k-1} \sum_{U \subset M: |U|=k} \mathbf{x}_U \cdot \dot{F}_0(\mathbf{x})|_{\mathbf{x}_U=\mathbf{1}} \right) d\mathbf{x} \right)^2.$$

Remark 3 Similar to $\hat{W}_{\emptyset,n}(\mathbf{x})$ one can get other empirical processes with a limiting covariance of the form (12) by subtracting from \mathbb{F}_n linear combinations of empirical processes of less dimension.

4.2 Local asymptotic optimality of independence tests

An interesting statistical problem that leads to the extremal problem (3)–(4) is that of local asymptotic optimality of independence tests.

Consider testing the hypothesis of independence $H_0 : \theta = 0$ versus the alternative $H_1 : \theta > 0$. Two commonly used measures for judging the quality of testing are the Pitman slope and the Bahadur local index. Under both approaches, the measure of efficiency of a given test statistic $T_n = T(\mathbf{X}_1, \dots, \mathbf{X}_n)$ has an upper bound (see, e.g., (van der Vaart, 1998, Th. 15.4) and Bahadur (1971)), which yields the inequality

$$b_T(\dot{F}_0) \leq \int_{I^m} \dot{f}_0^2(\mathbf{x}) d\mathbf{x}. \quad (21)$$

Here b_T is a homogeneous functional of degree 2 defined on the space \mathbf{H}^m that depends on a structure of T_n and measures efficiency of the corresponding test. For the Bahadur efficiency the upper bound (21) is a local version of the Bahadur–Ragavachari inequality (see Bahadur (1971)). For the test based on T_n , the closer $b_T(\dot{F}_0)$ is to $\int_{I^m} \dot{f}_0^2(\mathbf{x}) d\mathbf{x}$, the better the family $\{F_\theta : \theta \geq 0\}$ is. Thus, in order to describe the domain of Bahadur and/or Pitman optimality of the sequence of test statistics $\{T_n\}$, we need to know for which dependence function \dot{F}_0 equality in (21) is attained. If b_T is the square of a linear functional, this leads to extremal problem (3)–(4). For $m = 2$ some applications related to establishing Bahadur optimality of independence tests can be found in Nazarov and Nikitin (2000). Here we cite two examples from Nazarov and Nikitin (2000) with *non-Lebesgue* measure $\mu(\mathbf{x})$ in the problem (3)–(4) that corresponds to the integration over diagonal(s) of I^m . These examples are connected to testing independence using the Gini rank statistic and Spearman’s footrule.

Let $\mathbf{X}_i = (X_{i1}, \dots, X_{im})$, $m \geq 2$, $i = 1, \dots, n$, be as before. Denote by R_{ij} the rank of X_{ij} among X_{1j}, \dots, X_{nj} , $i = 1, \dots, n$, $j = 1, \dots, m$. Recall that the Gini rank coefficient is defined for $m = 2$ by

$$r_G = \frac{2}{D_n} \sum_{i=1}^n (|n+1 - R_{i1} - R_{i2}| - |R_{i1} - R_{i2}|),$$

where $D_n = n^2$ if n is even and $D_n = n^2 - 1$ if n is odd, and inequality (21) takes the form (see Nazarov and Nikitin (2000))

$$24 \left(\int_0^1 \left(\dot{F}_0(x, x) + \dot{F}_0(1-x, x) \right) dx \right)^2 \leq \int_{I^2} \dot{f}_0^2(\mathbf{x}) d\mathbf{x}. \quad (22)$$

In this case $\mu(\mathbf{x}) = \delta(x_1 - x_2) + \delta(1 - x_1 - x_2)$, $\mathbf{x} = (x_1, x_2) \in I^2$, and equality in (21) is attained for the function, cf. Theorem 1,

$$\dot{F}_0(\mathbf{x}) = C \int_{I^2} \mathcal{G}_{\mathcal{M}_0}(\mathbf{x}, \boldsymbol{\xi}) d\mu(\boldsymbol{\xi}), \quad C > 0, \quad (23)$$

where, in view of (14), $\mathcal{G}_{\mathcal{M}_\emptyset}(\mathbf{x}, \boldsymbol{\xi})$ is given by (17). Integrating in (23) yields

$$\dot{F}_0(\mathbf{x}) = C \left(|x_1 - x_2|^3 - |x_1 + x_2 - 1|^3 - 3(x_1^2 + x_2^2) + 3(x_1 + x_2 - 1) \right), \quad C > 0.$$

For the Spearman footrule based on the statistic $r_f = \sum_{i=1}^n |R_{i1} - R_{i2}|$ the local Bahadur index on the left-hand side of (21) equals

$$b_{r_f}(\dot{F}_0) = 90 \left(\int_0^1 \dot{F}_0(x, x) dx \right)^2,$$

which corresponds to the measure $\mu(\mathbf{x}) = \delta(x_1 - x_2)$, $\mathbf{x} = (x_1, x_2) \in I^2$. Therefore the optimal dependence function has the form

$$\begin{aligned} \dot{F}_0(\mathbf{x}) &= C \int_{I^2} \mathcal{G}_{\mathcal{M}_\emptyset}(\mathbf{x}, \boldsymbol{\xi}) d\mu(\boldsymbol{\xi}) \\ &= C \left(|x_1 - x_2|^3 - (x_1 + x_2)^3 + 2x_1x_2(x_1^2 + x_2^2 + 2) \right), \quad C > 0. \end{aligned}$$

Another interesting application, when $m \geq 2$, is connected to Pitman optimality of a multivariate Spearman's rho, cf. Schmid and Schmidt (2007); Quessy (2009):

$$S_{m,n} = \frac{1}{C_m} \left\{ n^{-1} \sum_{i=1}^n \prod_{k=1}^m (n+1 - R_{ik}) - \left(\frac{n+1}{2} \right)^m \right\},$$

where $C_m = n^{-1} \sum_{i=1}^n i^m - ((n+1)/2)^m$ is a normalizing factor. The statistic $S_{m,n}$ is a sample counterpart of the functional $s_m(F) = \frac{2^m(m+1)}{2^m - m - 1} \left(\int F dF_1 \dots dF_m - 2^{-m} \right)$.

As before, consider the family $\{F_\theta : \theta \geq 0\}$ of absolutely continuous copulas that satisfy **(A1)** and **(A2)**. Under the Pitman approach, $\theta = \theta_n = h/\sqrt{n}$, where $h \geq 0$ is a local parameter, and for all $h \geq 0$, cf. (Quessy, 2009, Cor. 1),

$$\frac{\sqrt{n}(S_{m,n} - \mu_m(h/\sqrt{n}))}{\sigma_m(0)} \overset{\mathbf{P}_{h/\sqrt{n}}^n}{\rightsquigarrow} \mathcal{N}(0, 1), \quad n \rightarrow \infty,$$

where $\mu_m(\theta) = \frac{2^m(m+1)}{2^m - m - 1} \theta \int_{I^m} \dot{F}_0(\mathbf{x}) d\mathbf{x}$ and $\sigma_m^2(0) = \frac{(m+1)^2((4/3)^m - m/3 - 1)}{(2^m - m - 1)^2}$. In view of Theorem 15.4 in van der Vaart (1998), inequality (21) takes the form

$$\frac{4^m}{(4/3)^m - m/3 - 1} \left(\int_{I^m} \dot{F}_0(\mathbf{x}) d\mathbf{x} \right)^2 \leq \int_{I^m} \dot{f}_0^2(\mathbf{x}) d\mathbf{x}. \quad (24)$$

Then, the application of Theorem 1 yields that the sequence of test statistics $\{S_{m,n}\}_{n \geq 1}$ is Pitman optimal if and only if, cf. (Stepanova, 2003, Ths. 2, 3) and (Quessy, 2009, Sec. 4.4),

$$\dot{F}_0(\mathbf{x}) = C \prod_{j=1}^m x_j \left(\prod_{j=1}^m (2 - x_j) + \sum_{j=1}^m x_j - (m+1) \right), \quad \mathbf{x} \in I^m, \quad C > 0. \quad (25)$$

Indeed, the test based on $S_{m,n}$ is the “best” for those dependence functions \dot{F}_0 that deliver equality in inequality (24). Thus, taking into account (14), we minimize the functional $\int_{I^m} \dot{f}_0^2(\mathbf{x}) d\mathbf{x}$ on the space \mathbf{H}^m subject to

$$\int_{I^m} \dot{F}_0(\mathbf{x}) d\mathbf{x} = 1, \quad \dot{F}_0(\mathbf{x})|_{\mathbf{x}_U=1} = 0, \quad \text{for any } U \in 2^M, \quad |U| = m - 1.$$

By the results of Sections 2 and 3, including Theorem 1, the functional $\int_{I^m} \dot{f}_0^2(\mathbf{x}) d\mathbf{x}$ is minimized when

$$\dot{F}_0(\mathbf{x}) = \lambda^{-1} \int_{I^m} \mathcal{G}_{\mathcal{M}_\emptyset}(\mathbf{x}, \boldsymbol{\xi}) d\boldsymbol{\xi}, \quad (26)$$

where $\mathcal{G}_{\mathcal{M}_\emptyset}$ is given by (17). By homogeneity of inequality (24) the extremal function is defined up to a positive constant. Integrating in (26) yields (25).

Remark 4 For $m = 2$ the test statistics $B_{\emptyset,n}^1$ and $S_{m,n}$ are known to be asymptotically equivalent (see, e.g., (Nikitin, 1995, Ch. 5)). The results of this section extend this property to all $m \geq 2$. Indeed, due to (19) and (24) the square of Pitman slope of $S_{m,n}$ equals the Bahadur local index of $B_{\emptyset,n}^1$. Thus, the respective left-hand sides in (21) coincide, and the tests based on $B_{\emptyset,n}^1$ and $S_{m,n}$, $m \geq 2$, are asymptotically equivalent.

Acknowledgments

The research of A. Nazarov was partly supported by RFBR grant 10-01-00154a. The research of N. Stepanova was supported by an NSERC grant. We would like to thank Prof. Yu. V. Tarannikov for communicating with us regarding references Kleitman (1969); Korshunov (1981). We are grateful to Prof. Ya. Yu. Nikitin for discussions and suggestions.

References

- Alexeev, V.M., Tikhomirov, V.M., and Fomin, S.V. (1979), *Optimal Control*, Moscow: Nauka (in Russian). English Translation, New York: Consultants Bureau, 1987.
- Bahadur, R.R. (1971), *Some Limit Theorems in Statistics*, Philadelphia: SIAM.
- Blum, J.R., Kiefer, J., and Rosenblatt, M. (1961), “Distribution Free Tests of Independence Based on the Sample Distribution Function”, *Annals of Mathematical Statistics*, 32, 485–498.
- Deheuvels P. (1981), “An Asymptotic Decomposition for Multivariate Distribution-Free Tests of Independence”, *Journal of Multivariate Analysis*, 11, 102–113.

- Deheuvels P. (2005), “Weighted Multivariate Cramér–von Mises-type Statistics”, *Afrika Statistika*, 1, 1–14.
- Dugué D. (1975), “Sur des tests d’indépendance indépendants de la loi”, *Comptes Rendus de l’Académie des Sciences, Sér. A*, 281, 1103–1104.
- Genest, C., Quessy, J.-F., and Rémillard B. (2007), “Asymptotic Local Efficiency of Cramér–von Mises Type Tests for Multivariate Dependence”, *Annals of Statistics*, 35, 166–191.
- Kleitman, D. (1969), “On Dedekind Problem: the Number of Monotone Boolean Functions”, *Proceedings of the American Mathematical Society*, 21, 677–682.
- Korshunov, A.D. (1981), “On Quantity of Monotone Functions”, *Problems of Cybernetics*, 38, pp. 5–108. (In Russian.)
- Nazarov, A.I., Nikitin, Ya.Yu. (2000), “Some Extremal Problems for Gaussian and Empirical Random Fields”, in *Proceedings of the Saint-Petersburg Mathematical Society*, ed. N.N. Uraltseva, 8, pp. 214–230 (in Russian). English translation in *Translations of the American Mathematical Society, Ser. 2*, 205 (2002), pp. 189–202.
- Neuhaus, G. (1971), “On Weak Convergence of Stochastic Processes with Multidimensional Time Parameter”, *Annals of Mathematical Statistics*, 42, 1285–1295.
- Nikitin Y. (1995), *Asymptotic Efficiency of Nonparametric Tests*, Cambridge University Press.
- Quessy, J.-F. (2009), “Theoretical Efficiency Comparisons of Independence Tests Based on Multivariate Versions of Spearman’s Rho”, *Metrika*, 70, 315–338.
- Schmid F., Schmidt R. (2007), “Multivariate Extensions of Spearman’s Rho and Related Statistics”, *Statistics and Probability Letters*, 77, 407–416.
- Stepanova, N.A. (2003), “Multivariate Rank Statistics for Independence and Their Asymptotic Efficiency”, *Mathematical Methods of Statistics*, 12, No. 2, 197–217.
- van der Vaart, A.W. (1998), *Asymptotic Statistics*, Cambridge University Press.